Gag_AF110965_BW_mod

ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGGCAAGCTGGACGCCTGGGAGCGCATCCGCC
TGCGCCCCGGCGGCAAGAAGTGCTACATGATGAAGCACCTGGTGTGGGCCAGCCGCGAGCT
GGAGAAGTTCGCCCTGAACCCCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATCATC
CGCCAGCTGCACCCCGCCCTGCAGACCGGCAGCGAGGAGCTGAAGAGCCTGTTCAACACCG
TGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTCCGCGACACCAAGGAGGCCCTGGA
CAAGATCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGATCCAGCAGGCCGAGGCCGCCGAC
AAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACC
AGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAGGCCTTCAG
CCCCGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAGGACCTGAAC
ACGATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGACACCATCA
ACGAGGAGGCCGCCGAGTGGGACCGCGTGCACCCCGTGCACGCCGGCCCCATCGCCCCCGG
CCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGCACCACCAGCACCCTGCAGGAGCAG
ATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTACAAGCGGTGGATCA
TCCTGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATCCTGGACATCAAGCA
GGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCCTGCGCGCCGAG
CAGAGCACCCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTGGTGCAGAACGCCAACC
CCGACTGCAAGACCATCCTGCGCGCTCTCGGCCCCGGCGCCAGCCTGGAGGAGATGATGAC
CGCCTGCCAGGGCGTGGGCGGCCCCAGCCACAAGGCCCGCGTGCTGGCCGAGGCGATGAGC
CAGGCCAACACCAGCGTGATGATGCAGAAGAGCAACTTCAAGGGCCCCCGGCGCATCGTCA
AGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAACTGCCGCGCCCCCCGCAAGAA
GGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCC
AACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGCCCCGGCAACTTCCTGCAGAGCC
GCCCCGAGCCCACCGCCCCCCCCGCCGAGAGCTTCCGCTTCGAGGAGACCACCCCCGGCCA
GAAGCAGGAGAGCAAGGACCGCGAGACCCTGACCAGCCTGAAGAGCCTGTTCGGCAACGAC
CCCCTGAGCCAGTAA

FIG. 1

Gag_AF110967_BW_mod

ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGAGAAGCTGGACAAGTGGGAGAAGATCCGCC
TGCGCCCCGGCGGCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAGCT
GGAGGGCTTCGCCCTGAACCCCGGCCTGCTGGAGACCGCCGAGGGCTGCAAGCAGATCATG
AAGCAGCTGCAGCCCGCCCTGCAGACCGGCACCGAGGAGCTGCGCAGCCTGTACAACACCG
TGGCCACCCTGTACTGCGTGCACGCCGGCATCGAGGTCCGCGACACCAAGGAGGCCCTGGA
CAAGATCGAGGAGGAGCAGAACAAGTCCCAGCAGAAGACCCAGCAGGCCAAGGAGGCCGAC
GGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCAGG
CCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAGGCCTTCAGCCC
CGAGGTGATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAGGACCTGAACACG
ATGTTGAACACCGTGGGCGGCCACCAGGCCGCCATGCAGATGCTGAAGGACACCATCAACG
AGGAGGCCGCCGAGTGGGACCGCCTGCACCCCGTGCAGGCCGGCCCCGTGGCCCCCGGCCA
GATGCGCGACCCCCGCGGCAGCGACATCGCCGGCGCCACCAGCACCCTGCAGGAGCAGATC
GCCTGGATGACCAGCAACCCCCCCGTGCCCGTGGGCGACATCTACAAGCGGTGGATCATCC
TGGGCCTGAACAAGATCGTGCGGATGTACAGCCCCGTGAGCATCCTGGACATCCGCCAGGG
CCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCCTGCGCGCCGAGCAG
GCCACCCAGGACGTGAAGAACTGGATGACCGAGACCCTGCTGGTGCAGAACGCCAACCCCG
ACTGCAAGACCATCCTGCGCGCTCTCGGCCCCGGCGCCACCCTGGAGGAGATGATGACCGC
CTGCCAGGGCGTGGGCGGCCCCGGCCACAAGGCCCGCGTGCTGGCCGAGGCGATGAGCCAG
GCCAACAGCGTGAACATCATGATGCAGAAGAGCAACTTCAAGGGCCCCCGGCGCAACGTCA
AGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAGAACTGCCGCGCCCCCCGCAAGAA
GGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCC
AACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGCCCCGGCAACTTCCTGCAGAACC
GCAGCGAGCCCGCCGCCCCCACCGTGCCCACCGCCCCCCCCGCCGAGAGCTTCCGCTTCGA
GGAGACCACCCCCGCCCCCAAGCAGGAGCCCAAGGACCGCGAGCCCTACCGCGAGCCCCTG
ACCGCCCTGCGCAGCCTGTTCGGCAGCGGCCCCCTGAGCCAGTAA

FIG. 2

Env_AF110968_C_BW_opt

--> signal peptide (1-81)

ATGCGCGTGATGGGCATCCTGAAGAACTACCAGCAGTGGTGGATGTGGGGCATCCTGGGCTTCTGGATGCTGATCA

\/--> gp120/140/160 (82)

TCAGCAGCGTGGTGGGCAACCTGTGGGTGACCGTGTACTACGGCGTGCCCGTGTGGAAGGAGGCCAAGACCACCCT

GTTCTGCACCAGCGACGCCAAGGCCTACGAGACCGAGGTGCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACC

GACCCCAACCCCCAGGAGATCGTGCTGGAGAACGTGACCGAGAACTTCAACATGTGGAAGAACGACATGGTGGACC

AGATGCACGAGGACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCTGCGTGAAGCTGACCCCCCTGTGCGTGAC

CCTGAAGTGCCGCAACGTGAACGCCACCAACAACATCAACAGCATGATCGACAACAGCAACAAGGGCGAGATGAAG

AACTGCAGCTTCAACGTGACCACCGAGCTGCGCGACCGCAAGCAGGAGGTGCACGCCCTGTTCTACCGCCTGGACG

TGGTGCCCCTGCAGGGCAACAACAGCAACGAGTACCGCCTGATCAACTGCAACACCAGCGCCATCACCCAGGCCTG

CCCCAAGGTGAGCTTCGACCCCATCCCCATCCACTACTGCACCCCCGCCGGCTACGCCATCCTGAAGTGCAACAAC

CAGACCTTCAACGGCACCGGCCCCTGCAACAACGTGAGCAGCGTGCAGTGCGCCCACGGCATCAAGCCCGTGGTGA

GCACCCAGCTGCTGCTGAACGGCAGCCTGGCCAAGGGCGAGATCATCATCCGCAGCGAGAACCTGGCCAACAACGC

CAAGATCATCATCGTGCAGCTGAACAAGCCCGTGAAGATCGTGTGCGTGCGCCCCAACAACAACACCCGCAAGAGC

GTGCGCATCGGCCCCGGCCAGACCTTCTACGCCACCGGCGAGATCATCGGCGACATCCGCCAGGCCTACTGCATCA

TCAACAAGACCGAGTGGAACAGCACCCTGCAGGGCGTGAGCAAGAAGCTGGAGGAGCACTTCAGCAAGAAGGCCAT

CAAGTTCGAGCCCAGCAGCGGCGGCGACCTGGAGATCACCACCCACAGCTTCAACTGCCGCGGCGAGTTCTTCTAC

TGCGACACCAGCCAGCTGTTCAACAGCACCTACAGCCCCAGCTTCAACGGCACCGAGAACAAGCTGAACGGCACCA

TCACCATCACCTGCCGCATCAAGCAGATCATCAACATGTGGCAGAAGGTGGGCCGCGCCATGTACGCCCCCCCCAT

CGCCGGCAACCTGACCTGCGAGAGCAACATCACCGGCCTGCTGCTGACCCGCGACGGCGGCAAGACCGGCCCCAAC

GACACCGAGATCTTCCGCCCCGGCGGCGGCGACATGCGCGACAACTGGCGCAACGAGCTGTACAAGTACAAGGTGG

gp120(1512)<--\/-->(1513)gp41

TGGAGATCAAGCCCCTGGGCGTGGCCCCCACCGAGGCCAAGCGCCGCGTGGTGGAGCGCGAGAAGCGCGCCGTGGG

CATCGGCGCCGTGTTCCTGGGCTTCCTGGGCGCCGCCGGCAGCACCATGGGCGCCGCCAGCATCACCCTGACCGTG

CAGGCCCGCCTGCTGCTGAGCGGCATCGTGCAGCAGCAGAACAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACC

TGCTGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGACCCGCATCCTGGCCGTGGAGCGCTACCTGAAGGACCA

GCAGCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCGCCGTGCCCTGGAACAGCAGCTGGAGC

AACCGCAGCCACGACGAGATCTGGGACAACATGACCTGGATGCAGTGGGACCGCGAGATCAACAACTACACCGACA

CCATCTACCGCCTGCTGGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGAAGGACCTGCTGGCCCTGGACAGCTG

gp140(2025)<--\/

GCAGAACCTGTGGAACTGGTTCAGCATCACCAACTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGC

CTGATCGGCCTGCGCATCATCTTCGCCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGCCCT

TCCAGACCCTGACCCCCAACCCCCGCGAGCCCGACCGCCTGGGCCGCATCGAGGAGGAGGGCGGCGAGCAGGACCG

CGGCCGCAGCATCCGCCTGGTGAGCGGCTTCCTGGCCCTGGCCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGC

TACCACCGCCTGCGCGACTTCATCCTGATCGCCGCCCGCGTGCTGGAGCTGCTGGGCCAGCGCGGCTGGGAGGCCC

TGAAGTACCTGGGCAGCCTGGTGCAGTACTGGGGCCTGGAGCTGAAGAAGAGCGCCATCAGCCTGCTGGACACCAT

CGCCATCGCCGTGGCCGAGGGCACCGACCGCATCATCGAGTTCATCCAGCGCATCTGCCGCGCCATCCGCAACATC

gp160, gp41(2547)<--\

CCCCGCCGCATCCGCCAGGGCTTCGAGGCCGCCCTGCAGTAA

FIG. 3

Env_AF110975_C_BW_opt

--> signal peptide (1-72) \/-->
ATGCGCGTGCGCGGCATCCTGCGCAGCTGGCAGCAGTGGTGGATCTGGGGCATCCTGGGCTTCTGGATCTGCAGCG
gp120/140/160 (72)
GCCTGGGCAACCTGTGGGTGACCGTGTACGACGGCGTGCCCGTGTGGCGCGAGGCCAGCACCACCCTGTTCTGCGC
CAGCGACGCCAAGGCCTACGAGAAGGAGGTGCACAACGTGTGGGCCACCCACGCCTGCGTGCCCACCGACCCCAAC
CCCCAGGAGATCGAGCTGGACAACGTGACCGAGAACTTCAACATGTGGAAGAACGACATGGTGGACCAGATGCACG
AGGACATCATCAGCCTGTGGGACCAGAGCCTGAAGCCCCGCGTGAAGCTGACCCCCCTGTGCGTGACCCTGAAGTG
CACCAACTACAGCACCAACTACAGCAACACCATGAACGCCACCAGCTACAACAACAACACCACCGAGGAGATCAAG
AACTGCACCTTCAACATGACCACCGAGCTGCGCGACAAGAAGCAGCAGGTGTACGCCCTGTTCTACAAGCTGGACA
TCGTGCCCCTGAACAGCAACAGCAGCGAGTACCGCCTGATCAACTGCAACACCAGCGCCATCACCCAGGCCTGCCC
CAAGGTGAGCTTCGACCCCATCCCCATCCACTACTGCGCCCCCGCCGGCTACGCCATCCTGAAGTGCAAGAACAAC
ACCAGCAACGGCACCGGCCCCTGCCAGAACGTGAGCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGAGCA
CCCCCCTGCTGCTGAACGGCAGCCTGGCCGAGGGCGGCGAGATCATCATCCGCAGCAAGAACCTGAGCAACAACGC
CTACACCATCATCGTGCACCTGAACGACAGCGTGGAGATCGTGTGCACCCGCCCCAACAACAACACCCGCAAGGGC
ATCCGCATCGGCCCCGGCCAGACCTTCTACGCCACCGAGAACATCATCGGCGACATCCGCCAGGCCCACTGCAACA
TCAGCGCCGGCGAGTGGAACAAGGCCGTGCAGCGCGTGAGCGCCAAGCTGCGCGAGCACTTCCCCAACAAGACCAT
CGAGTTCCAGCCCAGCAGCGGCGGCGACCTGGAGATCACCACCCACAGCTTCAACTGCCGCGGCGAGTTCTTCTAC
TGCAACACCAGCAAGCTGTTCAACAGCAGCTACAACGGCACCAGCTACCGCGGCACCGAGAGCAACAGCAGCATCA
TCACCCTGCCCTGCCGCATCAAGCAGATCATCGACATGTGGCAGAAGGTGGGCCGCGCCATCTACGCCCCCCCCAT
CGAGGGCAACATCACCTGCAGCAGCAGCATCACCGGCCTGCTGCTGGCCCGCGACGGCGGCCTGGACAACATCACC
ACCGAGATCTTCCGCCCCCAGGGCGGCGACATGAAGGACAACTGGCGCAACGAGCTGTACAAGTACAAGGTGGTGG
gp120(1509)<--\/-->(1510)gp41
AGATCAAGCCCCTGGGCGTGGCCCCCACCGAGGCCAAGCGCCGCGTGGTGGAGCGCGAGAAGCGCGCCGTGGGCAT
CGGCGCCGTGATCTTCGGCTTCCTGGGCGCCGCCGGCAGCAACATGGGCGCCGCCAGCATCACCCTGACCGCCCAG
GCCCGCCAGCTGCTGAGCGGCATCGTGCAGCAGCAGAGCAACCTGCTGCGCGCCATCGAGGCCCAGCAGCACATGC
TGCAGCTGACCGTGTGGGGCATCAAGCAGCTGCAGGCCCGCGTGCTGGCCATCGAGCGCTACCTGAAGGACCAGCA
GCTGCTGGGCATCTGGGGCTGCAGCGGCAAGCTGATCTGCACCACCACCGTGCCCTGGAACAGCAGCTGGAGCAAC
AAGACCCAGGGCGAGATCTGGGAGAACATGACCTGGATGCAGTGGGACAAGGAGATCAGCAACTACACCGGCATCA
TCTACCGCCTGCTGGAGGAGAGCCAGAACCAGCAGGAGCAGAACGAGAAGGACCTGCTGGCCCTGGACAGCCGCAA
gp140(2022)<--\/
CAACCTGTGGAGCTGGTTCAACATCAGCAACTGGCTGTGGTACATCAAGATCTTCATCATGATCGTGGGCGGCCTG
ATCGGCCTGCGCATCATCTTCGCCGTGCTGAGCATCGTGAACCGCGTGCGCCAGGGCTACAGCCCCCTGAGCTTCC
AGACCCTGACCCCCAACCCCCGCGGCCTGGACCGCCTGGGCCGCATCGAGGAGGAGGGCGGCGAGCAGGACCGCGA
CCGCAGCATCCGCCTGGTGCAGGGCTTCCTGGCCCTGGCCTGGGACGACCTGCGCAGCCTGTGCCTGTTCAGCTAC
CACCGCCTGCGCGACCTGATCCTGGTGACCGCCCGCGTGGTGGAGCTGCTGGGCCGCAGCAGCCCCCGCGGCCTGC
AGCGCGGCTGGGAGGCCCTGAAGTACCTGGGCAGCCTGGTGCAGTACTGGGGCCTGGAGCTGAAGAAGAGCGCCAC
CAGCCTGCTGGACAGCATCGCCATCGCCGTGGCCGAGGGCACCGACCGCATCATCGAGGTGATCCAGCGCATCTAC
gp160, gp41(2565)<--\
CGCGCCTTCTGCAACATCCCCCCGCCGCGTGCGCCAGGGCTTCGAGGCCGCCCTGCAGTAA

FIG. 4

Gag_AF110965_BW_opt

ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGGCAAGCTGGACGCCTGGGAGCGCATCCGCCTGCGCCCCGG

CGGCAAGAAGTGCTACATGATGAAGCACCTGGTGTGGGCCAGCCGCGAGCTGGAGAAGTTCGCCCTGAACC

CCGGCCTGCTGGAGACCAGCGAGGGCTGCAAGCAGATCATCCGCCAGCTGCACCCCGCCCTGCAGACCGGC

AGCGAGGAGCTGAAGAGCCTGTTCAACACCGTGGCCACCCTGTACTGCGTGCACGAGAAGATCGAGGTGCG
C

CGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGCAGAACAAGAGCCAGCAGAAGATCCAGCAGGCCG

AGGCCGCCGACAAGGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCAC

CAGGCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAGGCCTTCAGCCCCGAGGT

GATCCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAGGACCTGAACACCATGCTGAACACCGTGG
G T

GCGGCCACCAGGCCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCGTG

CACCCCGTGCACGCCGGCCCCATCGCCCCCGGCCAGATGCGCGAGCCCCGCGGCAGCGACATCGCCGGCAC

CACCAGCACCCTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCATCCCCGTGGGCGACATCTACA

AGCGCTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACAGCCCCGTGAGCATCCTGGACATCAAG
G G

CAGGGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCCTGCGCGCCGAGCAGAGCAC

CCAGGAGGTGAAGAACTGGATGACCGACACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCC

TGCGCGCCCTGGGCCCCGGCGCCAGCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCAGC
T C

CACAAGGCCCGCGTGCTGGCCGAGGCCATGAGCCAGGCCAACACCAGCGTGATGATGCAGAAGAGCAACTT
G

CAAGGGCCCCCGCCGCATCGTGAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCCGCAACTGCCGCG
G C

CCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAG

GCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGCCCCGGCAACTTCCTGCAGAGCCGCCCCGA

GCCCACCGCCCCCCCCGCCGAGAGCTTCCGCTTCGAGGAGACCACCCCCGGCCAGAAGCAGGAGAGCAAGG

ACCGCGAGACCCTGACCAGCCTGAAGAGCCTGTTCGGCAACGACCCCCTGAGCCAGTAA

FIG. 5

Gag_AF110967_BW_opt

ATGGGCGCCCGCGCCAGCATCCTGCGCGGCGAGAAGCTGGACAAGTGGGAGAAGATCCGCCTGCGCCCCGG

CGGCAAGAAGCACTACATGCTGAAGCACCTGGTGTGGGCCAGCCGCGAGCTGGAGGGCTTCGCCCTGAACC

CCGGCCTGCTGGAGACCGCCGAGGGCTGCAAGCAGATCATGAAGCAGCTGCAGCCCGCCCTGCAGACCGGC

ACCGAGGAGCTGCGCAGCCTGTACAACACCGTGGCCACCCTGTACTGCGTGCACGCCGGCATCGAGGTGCG
C

CGACACCAAGGAGGCCCTGGACAAGATCGAGGAGGAGCAGAACAAGAGCCAGCAGAAGACCCAGCAGGCCA
TC

AGGAGGCCGACGGCAAGGTGAGCCAGAACTACCCCATCGTGCAGAACCTGCAGGGCCAGATGGTGCACCAG

GCCATCAGCCCCCGCACCCTGAACGCCTGGGTGAAGGTGATCGAGGAGAAGGCCTTCAGCCCCGAGGTGAT

CCCCATGTTCACCGCCCTGAGCGAGGGCGCCACCCCCCAGGACCTGAACACCATGCTGAACACCGTGGGCG
G T

GCCACCAGGCCGCCATGCAGATGCTGAAGGACACCATCAACGAGGAGGCCGCCGAGTGGGACCGCCTGCAC

CCCGTGCAGGCCGGCCCCGTGGCCCCCGGCCAGATGCGCGACCCCCGCGGCAGCGACATCGCCGGCGCCAC

CAGCACCCTGCAGGAGCAGATCGCCTGGATGACCAGCAACCCCCCCGTGCCCGTGGGCGACATCTACAAGC

GCTGGATCATCCTGGGCCTGAACAAGATCGTGCGCATGTACAGCCCCGTGAGCATCCTGGACATCCGCCAG
G G

GGCCCCAAGGAGCCCTTCCGCGACTACGTGGACCGCTTCTTCAAGACCCTGCGCGCCGAGCAGGCCACCCA

GGACGTGAAGAACTGGATGACCGAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGACCATCCTGC

GCGCCCTGGGCCCCGGCGCCACCCTGGAGGAGATGATGACCGCCTGCCAGGGCGTGGGCGGCCCCGGCCAC
T C

AAGGCCCGCGTGCTGGCCGAGGCCATGAGCCAGGCCAACAGCGTGAACATCATGATGCAGAAGAGCAACTT
G

CAAGGGCCCCCGCCGCAACGTGAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCAAGAACTGCCGCG
G C

CCCCCCGCAAGAAGGGCTGCTGGAAGTGCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAG

GCCAACTTCCTGGGCAAGATCTGGCCCAGCCACAAGGGCCGCCCCGGCAACTTCCTGCAGAACCGCAGCGA

GCCCGCCGCCCCCACCGTGCCCACCGCCCCCCCCGCCGAGAGCTTCCGCTTCGAGGAGACCACCCCCGCCC

CCAAGCAGGAGCCCAAGGACCGCGAGCCCTACCGCGAGCCCCTGACCGCCCTGCGCAGCCTGTTCGGCAGC

GGCCCCCTGAGCCAGTAA

FIG. 6

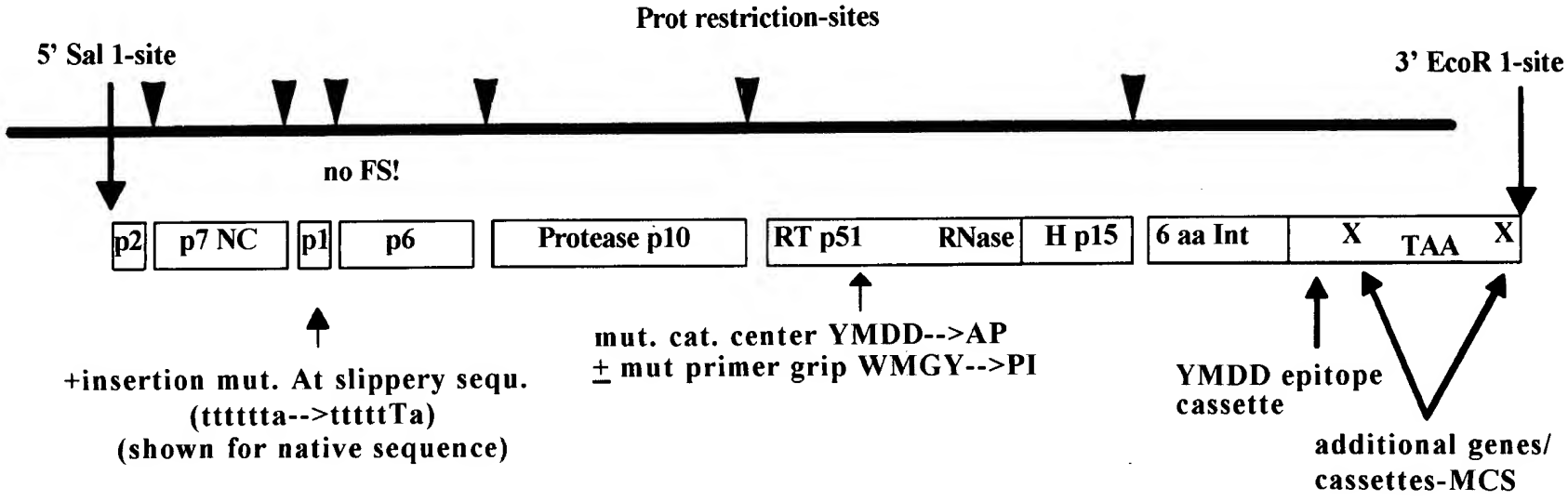

FIG. 7

PR975(+) (SEQ ID NO:30)

GTCGACGCCACCATGGCCGAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGAT
GCAGCGCAGCAACTTCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAA
GGAGGGCCACATCGCCCGCAACTGCCGCGCCCCCCCGCAAGAAGGGCTGCTGGAAGT
GCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTC
CGCGAGGACCTGGCCTTCCCCCAGGGCAAGGCCCGCGAGTTCCCCAGCGAGCAGAA
CCGCGCCAACAGCCCCACCAGCCGCGAGCTGCAGGTGCGCGGCGACAACCCCCGCA
GCGAGGCCGGCGCCGAGCGCCAGGGCACCCTGAACTTCCCCCAGATCACCCTGTGGC
AGCGCCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCAAGGAGGCCCTGCTGGAC
ACCGGCGCCGACGACACCGTGCTGGAGGAGATGAGCCTGCCCGGCAAGTGGAAGCC
CAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCT
GATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGATCGGCCCCACCCCCGT
GAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGCACCCTGAACTTCCCCAT
CAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGG
TGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGCGAG
GAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACAC
CCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACT
TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCC
ACCCCGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCC
TACTTCAGCGTGCCCCTGGACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCC
AGCATCAACAACGAGACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGC
TGGAAGGGCAGCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTC
CGCGCCCGCAACCCCGAGATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGC
AGCGACCTGGAGATCGGCCAGCACCGCGCCAAGATCGAGGAGCTGCGCAAGCACCT
GCTGCGCTGGGGCTTCACCACCCCCGACAAGAAGCACCAGAAGGAGCCCCCCTTCCT
GTGGATGGGCTACGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCGAGCTGCC
CGAGAAGGAGAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGAACT
GGGCCAGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGCTGCGCG
GCGCCAAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGCCGAGCTGGAGCTG
GCCGAGAACCGCGAGATCCTGCGCGAGCCCGTGCACGGCGTGTACTACGACCCCAG
CAAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCACGACCAGTGGACCTACCAGA
TCTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGCACC
GCCCACACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGA
GAGCATCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGAC
CTGGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTT
CGTGAACACCCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCAT
CGGCGCCGAGACCTTCTACGTGGACGGCGCCGCCAACCGCGAGACCAAGATCGGCA
AGGCCGGCTACGTGACCGACCGGGGCCGGCAGAAGATCGTGAGCCTGACCGAGACC
ACCAACCAGAAGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCAG
CGAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCC
CGACAAGAGCGAGAGCGAGCTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGG
AGAAGGTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCGGCGGCAACGAGCAG
ATCGACAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTGTTCCTGGACGGCATCGAT
GGCGGCATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGGCGGCCCT
AGGATCGATTAAAAGCTTCCCGGGGCTAGCACCGGTGAATTC

FIG. 8

PR975YM (SEQ ID NO:31)

GTCGACGCCACCATGGCCGAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGAT
GCAGCGCAGCAACTTCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAA
GGAGGGCCACATCGCCCGCAACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGT
GCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTC
CGCGAGGACCTGGCCTTCCCCCAGGGCAAGGCCCGCGAGTTCCCCAGCGAGCAGAA
CCGCGCCAACAGCCCCACCAGCCGCGAGCTGCAGGTGCGCGGCGACAACCCCCGCA
GCGAGGCCGGCGCCGAGCGCCAGGGCACCCTGAACTTCCCCCAGATCACCCTGTGGC
AGCGCCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCAAGGAGGCCCTGCTGGAC
ACCGGCGCCGACGACACCGTGCTGGAGGAGATGAGCCTGCCCGGCAAGTGGAAGCC
CAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCT
GATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGATCGGCCCCACCCCCGT
GAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGCACCCTGAACTTCCCCAT
CAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGG
TGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGCGAG
GAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACAC
CCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACT
TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCC
ACCCCGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCC
TACTTCAGCGTGCCCCTGGACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCC
AGCATCAACAACGAGACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGC
TGGAAGGGCAGCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTC
CGCGCCCGCAACCCCGAGATCGTGATCTACCAGGCCCCCCTGTACGTGGGCAGCGAC
CTGGAGATCGGCCAGCACCGCGCCAAGATCGAGGAGCTGCGCAAGCACCTGCTGCG
CTGGGGCTTCACCACCCCCGACAAGAAGCACCAGAAGGAGCCCCCCTTCCTGTGGAT
GGGCTACGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGA
AGGAGAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCC
AGCCAGATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGCTGCGCGGCGCC
AAGGCCCTGACCGACATCGTGCCCCTGACCGAGGAGGCCGAGCTGGAGCTGGCCGA
GAACCGCGAGATCCTGCGCGAGCCCGTGCACGGCGTGTACTACGACCCCAGCAAGG
ACCTGGTGGCCGAGATCCAGAAGCAGGGCCACGACCAGTGGACCTACCAGATCTAC
CAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGCACCGCCCA
CACCAACGACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAGAGCA
TCGTGATCTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGGG
AGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGA
ACACCCCCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCG
CCGAGACCTTCTACGTGGACGGCGCCGCCAACCGCGAGACCAAGATCGGCAAGGCC
GGCTACGTGACCGACCGGGGCCGGCAGAAGATCGTGAGCCTGACCGAGACCACCAA
CCAGAAGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCAGCGAGG
TGAACATCGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACA
AGAGCGAGAGCGAGCTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAG
GTGTACCTGAGCTGGGTGCCCGCCCACAAGGGCATCGGCGGCAACGAGCAGATCGA
CAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTGTTCCTGGACGGCATCGATGGCG
GCATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGGCGGCCCTAGGA
TCGATTAAAAGCTTCCCGGGGCTAGCACCGGTGAATTC

FIG. 9

PR975YMWM (SEQ ID NO:32)

GTCGACGCCACCATGGCCGAGGCCATGAGCCAGGCCACCAGCGCCAACATCCTGAT
GCAGCGCAGCAACTTCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAA
GGAGGGCCACATCGCCCGCAACTGCCGCGCCCCCCGCAAGAAGGGCTGCTGGAAGT
GCGGCAAGGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTC
CGCGAGGACCTGGCCTTCCCCCAGGGCAAGGCCCGCGAGTTCCCCAGCGAGCAGAA
CCGCGCCAACAGCCCCACCAGCCGCGAGCTGCAGGTGCGCGGCGACAACCCCCGCA
GCGAGGCCGGCGCCGAGCGCCAGGGCACCCTGAACTTCCCCCAGATCACCCTGTGGC
AGCGCCCCCTGGTGAGCATCAAGGTGGGCGGCCAGATCAAGGAGGCCCTGCTGGAC
ACCGGCGCCGACGACACCGTGCTGGAGGAGATGAGCCTGCCCGGCAAGTGGAAGCC
CAAGATGATCGGCGGCATCGGCGGCTTCATCAAGGTGCGCCAGTACGACCAGATCCT
GATCGAGATCTGCGGCAAGAAGGCCATCGGCACCGTGCTGATCGGCCCCACCCCCGT
GAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCTGCACCCTGAACTTCCCCAT
CAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGCCCCAAGG
TGAAGCAGTGGCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGCGAG
GAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCCGAGAACCCCTACAACAC
CCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACT
TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCC
ACCCCGCCGGCCTGAAGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGGCGACGCC
TACTTCAGCGTGCCCCTGGACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCC
AGCATCAACAACGAGACCCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGC
TGGAAGGGCAGCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTC
CGCGCCCGCAACCCCGAGATCGTGATCTACCAGGCCCCCCTGTACGTGGGCAGCGAC
CTGGAGATCGGCCAGCACCGCGCCAAGATCGAGGAGCTGCGCAAGCACCTGCTGCG
CTGGGGCTTCACCACCCCCGACAAGAAGCACCAGAAGGAGCCCCCCTTCCTGCCCAT
CGAGCTGCACCCCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGA
GCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAG
ATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGCTGCGCGGCGCCAAGGCC
CTGACCGACATCGTGCCCCTGACCGAGGAGGCCGAGCTGGAGCTGGCCGAGAACCG
CGAGATCCTGCGCGAGCCCGTGCACGGCGTGTACTACGACCCCAGCAAGGACCTGGT
GGCCGAGATCCAGAAGCAGGGCCACGACCAGTGGACCTACCAGATCTACCAGGAGC
CCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGCACCGCCCACACCAAC
GACGTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCGCCATGGAGAGCATCGTGAT
CTGGGGCAAGACCCCCAAGTTCCGCCTGCCCATCCAGAAGGAGACCTGGGAGACCT
GGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAACACCC
CCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGCCCATCATCGGCGCCGAG
ACCTTCTACGTGGACGGCGCCGCCAACCGCGAGACCAAGATCGGCAAGGCCGGCTA
CGTGACCGACCGGGGCCGGCAGAAGATCGTGAGCCTGACCGAGACCACCAACCAGA
AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCAGCGAGGTGAAC
ATCGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCAGCCCGACAAGAG
CGAGAGCGAGCTGGTGAACCAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGT
ACCTGAGCTGGGTGCCCGCCCACAAGGGCATCGGCGGCAACGAGCAGATCGACAAG
CTGGTGAGCAAGGGCATCCGCAAGGTGCTGTTCCTGGACGGCATCGATGGCGGCATC
GTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGGCGGCCCTAGGATCGAT
TAAAAGCTTCCCGGGGCTAGCACCGGTGAATTC

FIG. 10

8_5_ZA (SEQ ID NO:33)

```
1 TGGAAGGGTT AATTTACTCC AAGAAAAGGC AAGAAATCCT TGATTTGTGG GTCTATCACA
61 CACAAGGCTT CTTCCCTGAT TGGCAAAACT ACACACCGGG GCCAGGGGTC AGATATCCAC
121 TGACCTTTGG ATGGTGCTAC AAGCTAGTGC CAGTTGACCC AGGGGAGGTG GAAGAGGCCA
181 ACGGAGGAGA AGACAACTGT TTGCTACACC CTATGAGCCA ACATGGAGCA GAGGATGAAG
241 ATAGAGAAGT ATTAAAGTGG AAGTTTGACA GCCTCCTAGC ACGCAGACAC ATGGCCCGCG
301 AGCTACATCC GGAGTATTAC AAAGACTGCT GACACAGAAG GGACTTTCCG CCTGGGACTT
361 TCCACTGGGG CGTTCCGGGA GGTGTGGTCT GGGCGGGACT TGGGAGTGGT CAACCCTCAG
421 ATGCTGCATA TAAGCAGCTG CTTTTCGCCT GTACTGGGTC TCTCTCGGTA GACCAGATCT
481 GAGCCTGGGA GCCCTCTGGC TATCTAGGGA ACCCACTGCT TAAGCCTCAA TAAAGCTTGC
541 CTTGAGTGCT TTAAGTAGTG TGTGCCCATC TGTTGTGTGA CTCTGGTAAC TAGAGATCCC
601 TCAGACCCTT TGTGGTAGTG TGGAAAATCT CTAGCAGTGG CGCCCGAACA GGGACCAGAA
661 AGTGAAAGTG AGACCAGAGG AGATCTCTCG ACGCAGGACT CGGCTTGCTG AAGTGCACAC
721 GGCAAGAGGC GAGAGGGGCG GCTGGTGAGT ACGCCAATTT TACTTGACTA GCGGAGGCTA
781 GAAGGAGAGA GATGGGTGCG AGAGCGTCAA TATTAAGCGG CGGAAAATTA GATAAATGGG
841 AAAGAATTAG GTTAAGGCCA GGGGGAAAGA AACATTATAT GTTAAAACAT CTAGTATGGG
901 CAAGCAGGGA GCTGGAAAGA TTTGCACTTA ACCCTGGCCT GTTAGAAACA TCAGAAGGCT
961 GTAAACAAAT AATAAAACAG CTACAACCAG CTCTTCAGAC AGGAACAGAG GAACTTAGAT
1021 CATTATTCAA CACAGTAGCA ACTCTCTATT GTGTACATAA AGGGATAGAG GTACGAGACA
1081 CCAAGGAAGC CTTAGACAAG ATAGAGGAAG AACAAAACAA ATGTCAGCAA AAAGCACAAC
1141 AGGCAAAAGC AGCTGACGAA AAGGTCAGTC AAAATTATCC TATAGTACAG AATGCCCAAG
1201 GGCAAATGGT ACACCAAGCT ATATCACCTA GAACATTGAA TGCATGGATA AAAGTAATAG
1261 AGGAAAAGGC TTTCAATCCA GAGGAAATAC CCATGTTTAC AGCATTATCA GAAGGAGCCA
1321 CCCCACAAGA TTTAAACACA ATGTTAAATA CAGTGGGGGG ACATCAAGCA GCCATGCAAA
1381 TGTTAAAAGA TACCATCAAT GAGGAGGCTG CAGAATGGGA TAGGACACAT CCAGTACATG
1441 CAGGGCCTGT TGCACCAGGC CAGATGAGAG AACCAAGGGG AAGTGACATA GCAGGAACTA
1501 CTAGTACCCT TCAGGAACAA ATAGCATGGA TGACAAGTAA TCCACCTATT CCAGTAGAAG
1561 ACATCTATAA AAGATGGATA ATTCTGGGGT TAAATAAAAT AGTAAGAATG TATAGCCCTG
1621 TTAGCATTTT GGACATAAAA CAAGGGCCAA AAGAACCCTT TAGAGACTAT GTAGACCGGT
1681 TCTTTAAAAC CTTAAGAGCT GAACAAGCTA CACAAGATGT AAAGAATTGG ATGACAGACA
1741 CCTTGTTGGT CCAAAATGCG AACCCAGATT GTAAGACCAT TTTAAGAGCA TTAGGACCAG
1801 GGGCCTCATT AGAAGAAATG ATGACAGCAT GTCAGGGAGT GGGAGGACCT AGCCATAAAG
1861 CAAGAGTGTT GGCTGAGGCA ATGAGCCAAG CAAACAGTAA CATACTAGTG CAGAGAAGCA
1921 ATTTTAAAGG CTCTAACAGA ATTATTAAAT GTTTCAACTG TGGCAAAGTA GGGCACATAG
1981 CCAGAAATTG CAGGGCCCCT AGGAAAAAGG GCTGTTGGAA ATGTGGACAG GAAGGACACC
2041 AAATGAAAGA CTGTACTGAG AGGCAGGCTA ATTTTTTAGG GAAAATTTGG CCTTCCCACA
2101 AGGGGAGGCC AGGGAATTTC CTCCAGAACA GACCAGAGCC AACAGCCCCA CCAGCAGAAC
2161 CAACAGCCCC ACCAGCAGAG AGCTTCAGGT TCGAGGAGAC AACCCCCGTG CCGAGGAAGG
2221 AGAAAGAGAG GGAACCTTTA ACTTCCCTCA AATCACTCTT TGGCAGCGAC CCCTTGTCTC
2281 AATAAAAGTA GAGGGCCAGA TAAAGGAGGC TCTCTTAGAC ACAGGAGCAG ATGATACAGT
2341 ATTAGAAGAA ATAGATTTGC CAGGGAAATG GAAACCAAAA ATGATAGGGG GAATTGGAGG
2401 TTTTATCAAA GTAAGACAGT ATGATCAAAT ACTTATAGAA ATTTGTGGAA AAAAGGCTAT
2461 AGGTACAGTA TTAGTAGGGC CTACACCAGT CAACATAATT GGAAGAAATC TGTTAACTCA
2521 GCTTGGATGC ACACTAAATT TTCCAATTAG TCCTATTGAA ACTGTACCAG TAAAATTAAA
2581 ACCAGGAATG GATGGCCCAA AGGTCAAACA ATGGCCATTG ACAGAAGAAA AAATAAAAGC
2641 ATTAACAGCA ATTTGTGAGG AAATGGAGAA GGAAGGAAAA ATTACAAAAA TTGGGCCTGA
2701 TAATCCATAT AACACTCCAG TATTTGCCAT AAAAAAGAAG GACAGTACTA AGTGGAGAAA
2761 ATTAGTAGAT TTCAGGGAAC TCAATAAAAG AACTCAAGAC TTTTGGGAAG TTCAATTAGG
2821 AATACCACAC CCAGCAGGAT TAAAAAAGAA AAAATCAGTG ACAGTGCTAG ATGTGGGGGA
2881 TGCATATTTT TCAGTTCCTT TAGATGAAAG CTTCAGGAAA TATACTGCAT TCACCATACC
```

FIG. 11A

```
2941 TAGTATAAAC AATGAAACAC CAGGGATTAG ATATCAATAT AATGTGCTGC CACAGGGATG
3001 GAAAGGATCA CCAGCAATAT TCCAGAGTAG CATGACAAAA ATCTTAGAGC CCTTCAGAGC
3061 AAAAAATCCA GACATAGTTA TCTATCAATA TATGGATGAC TTGTATGTAG GATCTGACTT
3121 AGAAATAGGG CAACATAGAG CAAAAATAGA AGAGTTAAGG GAACATTTAT TGAAATGGGG
3181 ATTTACAACA CCAGACAAGA AACATCAAAA AGAACCCCCA TTTCTTTGGA TGGGGTATGA
3241 ACTCCATCCT GACAAATGGA CAGTACAACC TATACTGCTG CCAGAAAAGG ATAGTTGGAC
3301 TGTCAATGAT ATACAGAAGT TAGTGGGAAA ATTAAACTGG GCAAGTCAGA TTTACCCAGG
3361 GATTAAAGTA AGGCAACTCT GTAAACTCCT CAGGGGGGCC AAAGCACTAA CAGACATAGT
3421 ACCACTAACT GAAGAAGCAG AATTAGAATT GGCAGAGAAC AGGGAAATTT TAAGAGAACC
3481 AGTACATGGA GTATATTATG ATCCATCAAA AGACTTGATA GCTGAAATAC AGAAACAGGG
3541 GCATGAACAA TGGACATATC AAATTTATCA AGAACCATTT AAAAATCTGA AAACAGGGAA
3601 GTATGCAAAA ATGAGGACTA CCCACACTAA TGATGTAAAA CAGTTAACAG AGGCAGTGCA
3661 AAAAATAGCC ATGGAAAGCA TAGTAATATG GGGAAAGACT CCTAAATTTA GACTACCCAT
3721 CCAAAAAGAA ACATGGGAGA CATGGTGGAC AGACTATTGG CAAGCCACCT GGATCCCTGA
3781 GTGGGAGTTT GTTAATACCC CTCCCCTAGT AAAATTATGG TACCAACTAG AAAAAGATCC
3841 CATAGCAGGA GTAGAAACTT TCTATGTAGA TGGAGCAACT AATAGGGAAG CTAAAATAGG
3901 AAAAGCAGGG TATGTTACTG ACAGAGGAAG GCAGAAAATT GTTACTCTAA CTAACACAAC
3961 AAATCAGAAG ACTGAGTTAC AAGCAATTCA GCTAGCTCTG CAGGATTCAG GATCAGAAGT
4021 AAACATAGTA ACAGACTCAC AGTATGCATT AGGAATCATT CAAGCACAAC CAGATAAGAG
4081 TGACTCAGAG ATATTTAACC AAATAATAGA ACAGTTAATA AACAAGGAAA GAATCTACCT
4141 GTCATGGGTA CCAGCACATA AAGGAATTGG GGGAAATGAA CAAGTAGATA AATTAGTAAG
4201 TAAGGGAATT AGGAAAGTGT TGTTTCTAGA TGGAATAGAT AAAGCTCAAG AAGAGCATGA
4261 AAGGTACCAC AGCAATTGGA GAGCAATGGC TAATGAGTTT AATCTGCCAC CCATAGTAGC
4321 AAAAGAAATA GTAGCTAGCT GTGATAAATG TCAGCTAAAA GGGGAAGCCA TACATGGACA
4381 AGTCGACTGT AGTCCAGGGA TATGGCAATT AGATTGTACC CATTTAGAGG GAAAAATCAT
4441 CCTGGTAGCA GTCCATGTAG CTAGTGGCTA CATGGAAGCA GAGGTTATCC CAGCAGAAAC
4501 AGGACAAGAA ACAGCATATT TTATATTAAA ATTAGCAGGA AGATGGCCAG TCAAAGTAAT
4561 ACATACAGAC AATGGCAGTA ATTTTACCAG TACTGCAGTT AAGGCAGCCT GTTGGTGGGC
4621 AGGTATCCAA CAGGAATTTG GAATTCCCTA CAATCCCCAA AGTCAGGGAG TGGTAGAATC
4681 CATGAATAAA GAATTAAAGA AAATAATAGG ACAAGTAAGA GATCAAGCTG AGCACCTTAA
4741 GACAGCAGTA CAAATGGCAG TATTCATTCA CAATTTTAAA AGAAAAGGGG GAATTGGGGG
4801 GTACAGTGCA GGGGAAAGAA TAATAGACAT AATAGCAACA GACATACAAA CTAAAGAATT
4861 ACAAAAACAA ATTATAAGAA TTCAAAATTT TCGGGTTTAT TACAGAGACA GCAGAGACCC
4921 TATTTGGAAA GGACCAGCCG AACTACTCTG GAAAGGTGAA GGGGTAGTAG TAATAGAAGA
4981 TAAAGGTGAC ATAAAGGTAG TACCAAGGAG GAAAGCAAAA ATCATTAGAG ATTATGGAAA
5041 ACAGATGGCA GGTGCTGATT GTGTGGCAGG TGGACAGGAT GAAGATTAGA GCATGGAATA
5101 GTTTAGTAAA GCACCATATG TATATATCAA GGAGAGCTAG TGGATGGGTC TACAGACATC
5161 ATTTTGAAAG CAGACATCCA AAAGTAAGTT CAGAAGTACA TATCCCATTA GGGGATGCTA
5221 GATTAGTAAT AAAAACATAT TGGGGTTTGC AGACAGGAGA AAGAGATTGG CATTTGGGTC
5281 ATGGAGTCTC CATAGAATGG AGACTGAGAG AATACAGCAC ACAAGTAGAC CCTGACCTGG
5341 CAGACCAGCT AATTCACATG CATTATTTTG ATTGTTTTAC AGAATCTGCC ATAAGACAAG
5401 CCATATTAGG ACACATAGTT TTTCCTAGGT GTGACTATCA AGCAGGACAT AAGAAGGTAG
5461 GATCTCTGCA ATACTTGGCA CTGACAGCAT TGATAAAACC AAAAAAGAGA AAGCCACCTC
5521 TGCCTAGTGT TAGAAAATTA GTAGAGGATA GATGGAACGA CCCCCAGAAG ACCAGGGGCC
5581 GCAGAGGGAA CCATACAATG AATGGACACT AGAGATTCTA GAAGAACTCA AGCAGGAAGC
5641 TGTCAGACAC TTTCCTAGAC CATGGCTCCA TAGCTTAGGA CAATATATCT ATGAAACCTA
5701 TGGGGATACT TGGACGGGAG TTGAAGCTAT AATAAGAGTA CTGCAACAAC TACTGTTCAT
5761 TCATTTCAGA ATTGGATGCC AACATAGCAG AATAGGCATC TTGCGACAGA GAAGAGCAAG
5821 AAATGGAGCC AGTAGATCCT AAACTAAAGC CCTGGAACCA TCCAGGAAGC CAACCTAAAA
5881 CAGCTTGTAA TAATTGCTTT TGCAAACACT GTAGCTATCA TTGTCTAGTT TGCTTTCAGA
```

FIG. 11B

```
5941 CAAAAGGTTT AGGCATTTCC TATGGCAGGA AGAAGCGGAG ACAGCGACGA AGCGCTCCTC
6001 CAAGTGGTGA AGATCATCAA AATCCTCTAT CAAAGCAGTA AGTACACATA GTAGATGTAA
6061 TGGTAAGTTT AAGTTTATTT AAAGGAGTAG ATTATAGATT AGGAGTAGGA GCATTGATAG
6121 TAGCACTAAT CATAGCAATA ATAGTGTGGA CCATAGCATA TATAGAATAT AGGAAATTGG
6181 TAAGACAAAA GAAAATAGAC TGGTTAATTA AAAGAATTAG GGAAAGAGCA GAAGACAGTG
6241 GCAATGAGAG TGATGGGGAC ACAGAAGAAT TGTCAACAAT GGTGGATATG GGGCATCTTA
6301 GGCTTCTGGA TGCTAATGAT TTGTAACACG GAGGACTTGT GGGTCACAGT CTACTATGGG
6361 GTACCTGTGT GGAGAGAAGC AAAAACTACT CTATTCTGTG CATCAGATGC TAAAGCATAT
6421 GAGACAGAAG TGCATAATGT CTGGGCTACA CATGCTTGTG TACCCACAGA CCCCAACCCA
6481 CAAGAAATAG TTTTGGGAAA TGTAACAGAA AATTTTAATA TGTGGAAAAA TAACATGGCA
6541 GATCAGATGC ATGAGGATAT AATCAGTTTA TGGGATCAAA GCCTAAAGCC ATGTGTAAAG
6601 TTGACCCCAC TCTGTGTCAC TTTAAACTGT ACAGATACAA ATGTTACAGG TAATAGAACT
6661 GTTACAGGTA ATACAAATGA TACCAATATT GCAAATGCTA CATATAAGTA TGAAGAAATG
6721 AAAAATTGCT CTTTCAATGC AACCACAGAA TTAAGAGATA AGAAACATAA AGAGTATGCA
6781 CTCTTTTATA AACTTGATAT AGTACCACTT AATGAAAATA GTAACAACTT TACATATAGA
6841 TTAATAAATT GCAATACCTC AACCATAACA CAAGCCTGTC CAAAGGTCTC TTTTGACCCG
6901 ATTCCTATAC ATTACTGTGC TCCAGCTGAT TATGCGATTC TAAAGTGTAA TAATAAGACA
6961 TTCAATGGGA CAGGACCATG TTATAATGTC AGCACAGTAC AATGTACACA TGGAATTAAG
7021 CCAGTGGTAT CAACTCAACT ACTGTTAAAT GGTAGTCTAG CAGAAGAAGG GATAATAATT
7081 AGATCTGAAA ATTTGACAGA GAATACCAAA ACAATAATAG TACATCTTAA TGAATCTGTA
7141 GAGATTAATT GTACAAGGCC CAACAATAAT ACAAGGAAAA GTGTAAGGAT AGGACCAGGA
7201 CAAGCATTCT ATGCAACAAA TGACGTAATA GGAAACATAA GACAAGCACA TTGTAACATT
7261 AGTACAGATA GATGGAATAA AACTTTACAA CAGGTAATGA AAAAATTAGG AGAGCATTTC
7321 CCTAATAAAA CAATAAAATT TGAACCACAT GCAGGAGGGG ATCTAGAAAT TACAATGCAT
7381 AGCTTTAATT GTAGAGGAGA ATTTTTCTAT TGCAATACAT CAAACCTGTT TAATAGTACA
7441 TACTACCCTA AGAATGGTAC ATACAAATAC AATGGTAATT CAAGCTTACC CATCACACTC
7501 CAATGCAAAA TAAAACAAAT TGTACGCATG TGGCAAGGGG TAGGACAAGC AATGTATGCC
7561 CCTCCCATTG CAGGAAACAT AACATGTAGA TCAAACATCA CAGGAATACT ATTGACACGT
7621 GATGGGGGAT TTAACAACAC AAACAACGAC ACAGAGGAGA CATTCAGACC TGGAGGAGGA
7681 GATATGAGGG ATAACTGGAG AAGTGAATTA TATAAATATA AAGTGGTAGA AATTAAGCCA
7741 TTGGGAATAG CACCCACTAA GGCAAAAAGA AGAGTGGTGC AGAGAAAAAA AAGAGCAGTG
7801 GGAATAGGAG CTGTGTTCCT TGGGTTCTTG GGAGCAGCAG GAAGCACTAT GGGCGCAGCG
7861 TCAATAACGC TGACGGTACA GGCCAGACAA CTGTTGTCTG GTATAGTGCA ACAGCAAAGC
7921 AATTTGCTGA AGGCTATAGA GGCGCAACAG CATATGTTGC AACTCACAGT CTGGGGCATT
7981 AAGCAGCTCC AGGCGAGAGT CCTGGCTATA GAAAGATACC TAAAGGATCA ACAGCTCCTA
8041 GGGATTTGGG GCTGCTCTGG AAGACTCATC TGCACCACTG CTGTGCCTTG GAACTCCAGT
8101 TGGAGTAATA AATCTGAAGC AGATATTTGG GATAACATGA CTTGGATGCA GTGGGATAGA
8161 GAAATTAATA ATTACACAGA AACAATATTC AGGTTGCTTG AAGACTCGCA AAACCAGCAG
8221 GAAAAGAATG AAAAAGATTT ATTAGAATTG GACAAGTGGA ATAATCTGTG GAATTGGTTT
8281 GACATATCAA ACTGGCTGTG GTATATAAAA ATATTCATAA TGATAGTAGG AGGCTTGATA
8341 GGTTTAAGAA TAATTTTTGC TGTGCTCTCT ATAGTGAATA GAGTTAGGCA GGGATACTCA
8401 CCTTTGTCAT TTCAGACCCT TACCCCAAGC CCGAGGGGAC TCGACAGGCT CGGAGGAATC
8461 GAAGAAGAAG GTGGAGAGCA AGACAGAGAC AGATCCATAC GATTGGTGAG CGGATTCTTG
8521 TCGCTTGCCT GGGACGATCT GCGGAGCCTG TGCCTCTTCA GCTACCACCG CTTGAGAGAC
8581 TTCATATTAA TTGCAGTGAG GGCAGTGGAA CTTCTGGGAC ACAGCAGTCT CAGGGGACTA
8641 CAGAGGGGGT GGGAGATCCT TAAGTATCTG GGAAGTCTTG TGCAGTATTG GGGTCTAGAG
8701 CTAAAAAAGA GTGCTATTAG TCCGCTTGAT ACCATAGCAA TAGCAGTAGC TGAAGGAACA
8761 GATAGGATTA TAGAATTGGT ACAAAGAATT TGTAGAGCTA TCCTCAACAT ACCTAGGAGA
8821 ATAAGACAGG GCTTTGAAGC AGCTTTGCTA TAAAATGGGA GGCAAGTGGT CAAAACGCAG
8881 CATAGTTGGA TGGCCTGCAG TAAGAGAAAG AATGAGAAGA ACTGAGCCAG CAGCAGAGGG
8941 AGTAGGAGCA GCGTCTCAAG ACTTAGATAG ACATGGGGCA CTTACAAGCA GCAACACACC
```

FIG. 11C

```
9001 TGCTACTAAT GAAGCTTGTG CCTGGCTGCA AGCACAAGAG GAGGACGGAG ATGTAGGCTT
9061 TCCAGTCAGA CCTCAGGTAC CTTTAAGACC AATGACTTAT AAGAGTGCAG TAGATCTCAG
9121 CTTCTTTTTA AAAGAAAAGG GGGGACTGGA AGGGTTAATT TACTCTAGGA AAAGGCAAGA
9181 AATCCTTGAT TTGTGGGTCT ATAACACACA AGGCTTCTTC CCTGATTGGC AAAACTACAC
9241 ATCGGGGCCA GGGGTCCGAT TCCCACTGAC CTTTGGATGG TGCTTCAAGC TAGTACCAGT
9301 TGACCCAAGG GAGGTGAAAG AGGCCAATGA AGGAGAAGAC AACTGTTTGC TACACCCTAT
9361 GAGCCAACAT GGAGCAGAGG ATGAAGATAG AGAAGTATTA AAGTGGAAGT TTGACAGCCT
9421 TCTAGCACAC AGACACATGG CCCGCGAGCT ACATCCGGAG TATTACAAAG ACTGCTGACA
9481 CAGAAGGGAC TTTCCGCCTG GGACTTTCCA CTGGGGCGTT CCGGGAGGTG TGGTCTGGGC
9541 GGGACTTGGG AGTGGTCACC CTCAGATGCT GCATATAAGC AGCTGCTTTT CGCTTGTACT
9601 GGGTCTCTCT CGGTAGACCA GATCTGAGCC TGGGAGCTCT CTGGCTATCT AGGGAACCCA
9661 CTGCTTAGGC CTCAATAAAG CTTGCCTTGA GTGCTCTAAG TAGTGTGTGC CCATCTGTTG
9721 TGTGACTCTG GTAACTAGAG ATCCCTCAGA CCCTTTGTGG TAGTGTGGAA AATCTCTAGC
9781 A
```

FIG. 11D

SEQ ID NO:34

GCTGAGGCAATGAGCCAAGCAACCAGCGCAAACATACTGATGCAGAGAAGCAATTT
CAAAGGCCCTAAAAGAATTATTAAATGTTTCAACTGTGGCAAGGAAGGGCACATAG
CTAGAAATTGTAGGGCCCCTAGGAAAAAAGGCTGTTGGAAATGTGGAAAGGAAGGA
CACCAAATGAAAGACTGTACTGAGAGGCAGGCTAA

FIG. 12

975Pol wt until 6aa Int: (SEQ ID NO:35)

TTTTTTAGGGAAGATTTGGCCTTCCCACAAGGGAAGGCCAGGGAATTTCCTTCAGAA
CAGAACAGAGCCAACAGCCCCACCAGCAGAGAGCTTCAAGTTCGAGGAGACAACCC
CCGCTCCGAAGCAGGAGCCGAAAGACAGGGAACCCTTAATTTCCCTCAAATCACTCT
TTGGCAGCGACCCCTTGTCTCAATAAAAGTAGGGGGTCAAATAAAGGAGGCTCTCTT
AGACACAGGAGCTGATGATACAGTATTAGAAGAAATGAGTTTGCCAGGAAAATGGA
AACCAAAAATGATAGGAGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCAA
ATACTTATAGAAATTTGTGGAAAAAAGGCTATAGGTACAGTATTAATAGGACCTACA
CCTGTCAACATAATTGGAAGGAATATGTTGACTCAGCTTGGATGCACACTAAATTTT
CCAATTAGTCCCATTGAAACTGTGCCAGTAAAATTAAAGCCAGGAATGGATGGCCCA
AAGGTTAAACAATGGCCATTGACAGAAGAGAAAATAAAAGCATTAACAGCAATTTG
TGAAGAAATGGAGAAAGAAGGAAAAATTACAAAAATTGGGCCTGAAAATCCATATA
ACACTCCAGTATTTGCCATAAAAAAGAAGGACAGTACTAAGTGGAGAAAGTTAGTA
GATTTCAGGGAACTTAATAAAAGAACTCAAGACTTTTGGGAAGTTCAATTAGGAATA
CCACACCCAGCAGGGTTAAAAAAGAAAAAATCAGTGACAGTACTGGATGTGGGGGA
TGCATATTTTTCAGTTCCTTTAGATGAGGACTTCAGGAAATATACTGCATTCACCATA
CCTAGTATAAACAATGAAACACCAGGGATTAGATATCAATATAATGTGCTTCCACAG
GGATGGAAAGGATCACCATCAATATTCCAGAGTAGCATGACAAAAATCTTAGAGCC
CTTTAGAGCAAGAAATCCAGAAATAGTCATCTATCAATATATGGATGACTTGTATGT
AGGATCTGACTTAGAAATAGGGCAACATAGAGCAAAAATAGAGGAGTTAAGAAAAC
ATCTGTTAAGGTGGGGATTTACCACACCGGACAAGAAACATCAGAAAGAACCCCCA
TTTCTTTGGATGGGGTATGAACTCCATCCTGACAAATGGACAGTACAGCCTATAGAG
TTGCCAGAAAAGGAAAGCTGGACTGTCAATGATATACAGAAGTTAGTGGGAAAATT
AAATTGGGCCAGTCAGATTTACCCAGGAATTAAAGTAAGGCAACTTTGTAAACTCCT
TAGGGGGGCCAAAGCACTAACAGATATAGTACCACTAACTGAAGAAGCAGAATTAG
AATTGGCAGAGAACAGGGAAATTCTAAGAGAACCAGTACATGGAGTATATTATGAC
CCATCAAAAGACTTGGTAGCTGAAATACAGAAACAGGGGCATGACCAATGGACATA
TCAAATTTACCAAGAACCATTCAAAAACCTGAAAACAGGGAAGTATGCAAAAATGA
GGACTGCCCACACTAATGATGTAAAACAGTTAACAGAGGCAGTGCAAAAAATAGCT
ATGGAAAGCATAGTAATATGGGGAAAGACTCCTAAATTTAGACTACCCATCCAAAA
AGAAACATGGGAGACATGGTGGACAGACTATTGGCAAGCCACCTGGATTCCTGAGT
GGGAGTTTGTTAATACCCCTCCCTTAGTAAAATTATGGTACCAGCTAGAGAAAGAAC
CCATAATAGGAGCAGAAACTTTCTATGTAGATGGAGCAGCTAATAGGGAAACTAAA
ATAGGAAAAGCAGGGTATGTTACTGACAGAGGAAGGCAGAAAATTGTTTCTCTAAC
AGAAACAACAAATCAGAAGACTGAATTACAAGCAATTCAGCTAGCTTTGCAAGATTC
AGGATCAGAAGTAAACATAGTAACAGACTCACAGTATGCATTAGGAATCATTCAAG
CACAACCAGATAAGAGTGAATCAGAGTTAGTCAACCAAATAATAGAACAATTAATA
AAAAAGGAAAAGGTCTACCTGTCATGGGTACCAGCACATAAAGGAATTGGAGGAAA
TGAACAAATAGATAAATTAGTAAGTAAGGGAATCAGGAAAGTGCTGTTTCTAGATG
GAATAGAT

FIG. 13

SEQ ID NO:36

GGCGGCATCGTGATCTACCAGTACATGGACGACCTGTACGTGGGCAGCGGCG
GC

FIG. 14

SEQ ID NO: 37

GGIVIYQYMDDLYVGSGG

FIG. 15

12_5/1ZA (SEQ ID NO:45)

TGGAAGGGTTAATTTACTCCAGGAAAAGGCAAGAGATCCTTGATTTATGGGTCTATC
ACACACAAGGCTACTTCCCTGATTGGCAAAACTACACACCGGGACCAGGGGTCAGA
TATCCACTGACCTTTGGATGGTGCTTCAAGCTAGTGCCAGTTGACCCAAGGGAAGTA
GAAGAGGCCAACGGAGGAGAAGACAACTGTTTGCTACACCCTATGAGCCAGTATGG
AATGGATGATGAACACAAAGAAGTGTTACAGTGGAAGTTTGACAGCAGCCTAGCAC
GCAGACACCTGGCCCGCGAGCTACATCCGGATTATTACAAAGACTGCTGACACAGA
AGGGACTTTCCGCCTGGGACTTTCCACTGGGGCGTTCCAGGGGGAGTGGTCTGGGCG
GGACTGGGAGTGGCCAGCCCTCAGATGCTGCATATAAGCAGCGGCTTTTCGCCTGTA
CTGGGTCTCTCTAGGTAGACCAGATCCGAGCCTGGGAGCTCTCTGTCTATCTGGGGA
ACCCACTGCTTAGGCCTCAATAAAGCTTGCCTTGAGTGCTCTAAGTAGTGTGTGCCC
ATCTGTTGTGTGACTCTGGTAACTCTGGTAACTAGAGATCCCTCAGACCCTTTGTGGT
AGTGTGGAAAATCTCTAGCAGTGGCGCCCGAACAGGGACTTGAAAGCGAAAGTGAG
ACCAGAGAAGATCTCTCGACGCAGGACTCGGCTTGCTGAAGTGCACTCGGCAAGAG
GCGAGGGGGGCGACTGGTGAGTACGCCAAAATTTTTTTTGACTAGCGGAGGCTAGA
AGGAGAGAGATGGGTGCGAGAGCGTCAATATTAAGAGGGGGAAAATTAGACAAAT
GGGAAAAAATTAGGTTACGGCCAGGGGGGAGAAAACACTATATGCTAAAACACCTA
GTATGGGCAAGCAGAGAGCTGGAAAGATTTGCAGTTAACCCTGGCCTTTTAGAGAC
ATCAGACGGATGTAGAC AAATAATAAAACAGCTACAACCAGCTCTTCAGA
CAGGAACAGAGGAAATTAGATCATTATTTAACACAGTAGCAACTCTCTATTGTGTAC
ATAAAGGGATAGATGTACGAGACACCAAGGAAGCCTTAGACAAGATAGAGGAGGA
ACAAAACAAATGTCAGCAAAAAACACAGCAGGCGGAAGCGGCTGACAAAAAGGTC
AGTCAAAATTATCCTATAGTGCAGAACCTCCAAGGGCAAATGGTACACCAGGCCAT
ATCACCTAGAACCTTGAATGCATGGGTAAAAGTAATAGAGGAGAAGGCTTTTAGCC
CAGAGGTAATACCCATGTTTACAGCATTATCAGAAGGAGCCACCCCACAAGATTTA
AACACCATGTTAAATACAGTGGGGGGACATCAAGCAGCCATGCAAATGTTAAAAG
ATACCATCAATGAGGAGGCTGCAGAATGGGATAGGTTACATCCAGTACATGCAGGG
CCTGTTGCACCAGGCCAGATGAGAGAACCAAGGGGAAGTGACATAGCAGGAACTA
CTAGTACCCTTCAAGAACAAATAGCATGGATGACAAGTAACCCACCTATCCCAGTA
GGGGACATCTATAAAAGGTGGATAATTCTGGGGTTAAATAAAATAGTAAGAATGTA
CAGCCCTGTCAGCATTTTAGACATAAAACAAGGACCAAAGGAACCCTTTAGAGACT
ATGTAGACCGGTTCTTCAAAACTTTAAGAGCTGAACAATCTACACAAGAGGTAAAA
AATTGGATGACAGACACCTTGTTAGTCCAAAATGCGAACCCAGATTGTAAGACCATT
TTAAGAGCATTAGGACCAGGGGCTTCATTAGAAGAAATGATGACAGCATGTCAGGG
AGTGGGAGGACCTAGCCACAAAGCAAGAGTTTTGGCTGAGGCAATGAGCCAAGCAA
ACAATACAAGTGTAATGATACAGAAAAGCAATTTTAAAGGCCCTAGAAGAGCTGTT
AAATGTTTCAACTGTGGCAGGGAAGGGCACATAGCCAGGAATTGCAGGGCCCCTAG
GAAAAGGGGCTGTTGGAAATGTGGAAAGGAAGGACACCAAATGAAAGACTGTACT
GAGAGGCAGGCTAATTTTTTAGGGAAAATTTGGCCTTCCCACAAGGGGAGGCCAGG
GAATTTCCTTCAGAGCAGACCAGAGCCAACAGCCCCACCACTAGAACCAACAGCCC
CACCAGCAGAGAGCTTCAAGTTCAAGGAGACTCCGAAGCAGGAGCCGAAAGACAG
GGAACCTTTAACTTCCCTCAAATCACTCTTTGGCAGCGACCCCTTGTCTCAATAAAA

FIG. 16A

GTAGCGGGCCAAACAAAGGAGGCTCTTTTAGATACAGGAGCAGATGATACAGTACT
AGAAGAAATAAACTTGCCAGGAAAATGGAAACCAAAAATGATAGGAGGAATTGGA
GGTTTTATCAAAGTAAGACAGTATGATCAAATACTTATAGAAATTTGTGGAAAAAGG
GCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTG
TTGACTCAGCTTGGATGCACACTAAATTTTCCAATTAGCCCCATTGAAACTGTACCA
GTAAAATTAAAGCCAGGAATGGATGGCCCAAAGGTTAAACAATGGCCATTGACAGA
AGAAAAAATAAAAGCATTAACAGAAATTTGTGAGGAAATGGAGAAGGAAGGAAAA
ATTACAAAAATTGGGCCTGAAAATCCATATAACACTCCAGTATTTGCCATAAAGAAG
AAGGACAGTACAAAGTGGAGAAAATTAGTAGATTTCAGGGAACTCAATAAAAGAAC
TCAAGACTTTTGGGAAGTCCAATTAGGAATACCACACCCAGCAGGGTTAAAAAAGA
AAAAATCAGTGACAGTACTGGATGTGGGAGATGCATATTTTTCAGTCCCTTTAGATG
AGAGCTTCAGAAAATATACTGCATTCACCATACCTAGTATAAACAATGAAACACCA
GGGATTAGATATCAATATAATGTTCTTCCACAGGGATGGAAAGGATCACCAGCAA
TATTCCAGAGTAGCATGACAAGAATCTTAGAGCCCTTTAGAACACAAAACCCAGAA
GTAGTTATCTATCAATATATGGATGACTTATATGTAGGATCTGACTTAGAAATAGGG
CAACATAGAGCAAAAATAGAGGAGTTAAGAGGACACCTATTGAAATGGGGATTTAC
CACACCAGACAAGAAACATCAGAAAGAACCCCCATTTCTTTGGATGGGGTATGAAC
TCCATCCTGACAAATGGACAGTACAGCCTATACAGCTGCCAGAAAAGGAGAGCTGG
ACTGTCAATGATATACAGAAGTTAGTGGGAAAGTTAAACTGGGCAAGTCAGATTTA
CCCAGGGATTAAAGTAAGGCAACTGTGTAAACTCCTTAGGGGAGCCAAAGCACTAA
CAGACATAGTGCCACTGACTGAAGAAGCAGAATTAGAATTGGCTGAGAACAGGGA
AATTCTAAAAGAACCAGTACATGGAGTATATTATGACCCATCAAAAGATTTAATAG
CTGAAATACAGAAACAGGGGAATGACCAATGGACATATCAAATTTACCAAGAACC
ATTTAAAAATCTGAGAACAGGAAAGTATGCAAAAATGAGGACTGCCCACACTAATG
ATGTGAAACAGTTAGCAGAGGCAGTGCAAAAGATAACCCAGGAAAGCATAGTAATA
TGGGGAAAAACTCCTAAATTTAGACTACCCATCCCAAAAGAAACATGGGAGACATG
GTGGTCAGACTATTGGCAAGCCACCTGGATTCCTGAGTGGGAGTTTGTCAATACCCC
TCCCCTAGTAAAATTGTGGTACCAGCTGGAAAAAGAACCCATAGTAGGGGCAGAAA
CTTTCTATGTAGATGGAGCAGCCAATAGGGAAACTAAAATAGGAAAAGCAGGGTAT
GTCACTGACAAAGGAAGGCAGAAAGTTGTTTCCTTCACTGAAACAACAAATCAGAA
GACTGAATTACAAGCAATTCAGCTAGCTTTGCAGGATTCAGGGCCAGAAGTAAACA
TAGTAACAGACTCACAGTATGCATTAGGAATCATTCAAGCACAACCAGATAAGAGT
GAATCAGAATTAGTCAGTCAAATAATAGAACAGTTGATAAAAAAGGAAAAAGTCTA
CCTATCATGGGTACCAGCACATAAAGGAATTGGAGGAAATGAACAAGTAGACAAAT
TAGTAAGTAGTGGAATCAGAAAAGTACTGTTTCTAGATGGAATAGATAAAGCTCAA
GAAGAGCATGAAAAATATCACAGCAATTGGAGAGCAATGGCTAGTGAGTTTAATCT
GCCACCCATAGTAGCAAAGGAAATAGTAGCCAGCTGTGATAAATGTCAGCTAAAAG
GGGAAGCCATGCATGGACAAGTCGACTGTAGTCCAGGAATATGGCAATTAGACTGT
ACACATTTAGAAGGAAAAATCATCCTAGTAGCAGTCCATGTAGCCAGTGGCTACAT
GGAAGCAGAGGTTATCCCAGCAGAAACAGGACAAGAAACAGCATACTTTATACTAA
AATTAGCAGGAAGATGGCCAGTCAAAGTAATACATACAGATAATGGCAGTAATTTC
ACCAGTACCGCAGTTAAGGCAGCCTGTTGGTGGGCAGATATCCAACGGGAATTTGG
AATTCCCTACAATCCCCAAAGTCAAGGAGTAGTAGAATCCATGAATAAAGAATTAA

FIG. 16B

AGAAAATCATAGGGCAAGTAAGAGATCAAGCTGAGCACCTTAAGACAGCAGTACAA
ATGGCAGTATTCATTCACAATTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGC
AGGGGAGAGAATAATAGACATAATAGCATCAGACATACAAACTAAAGAATTACAAA
AACAAATTATAAAAATTCAAAATTTTCGGGTTTATTACAGAGACAGCAGAGACCCTA
TTTGGAAAGGACCAGCCAAACTACTCTGGAAAGGTGAAGGGGCAGTAGTAATACAA
GATAATAGTGATATAAAGGTAGTACCAAGAAGGAAAGCAAAAATCATTAAGGACTA
TGGAAAACAGATGGCAGGTGCTGATTGTGTGGCAGGTAGACAGGATGAAGATTAGA
ACATGGCACAGTTTAGTAAAGCACCATATGTATGTTTCGAGGAGAGCTGATGGATGG
TTCTACAGACATCATTATGAAAGCAGACACCCAAAAGTAAGTTCAGAAGTACACAT
CCCATTAGGAGATGCCAGGTTAGTAATAAAAACATATTGGGGTCTGCAGACAGGAG
AAAGAGCTTGGCATTTGGGTCACGGAGTCTCCATAGAATGGAGATTGAGAAGATAT
AGCACACAAGTAGACCCTGACCTGACAGACCAACTAATTCATATGCATTATTTTGAT
TGTTTTGCAGAATCTGCCATAAGGAAAGCCATACTAGGACAGATAGTTAGCCCTAA
GTGTGACTATCAAGCAGGACATAACAAGGTAGGATCTCTACAATACTTGGCACTGA
CAGCATTGATAAAACCAAAAAAGATAAAGCCACCTCTGCCTAGTGTTAGGAAATTA
GTAGAGGATAGATGGAACAAGCCCCAGAAGACCAGGGGCCGCAGAGGGAACCATA
CAATGAATGGACACTAGAGCTTTTAGAAGAACTCAAGCAGGAAGCTGTCAGACACT
TTCCTAGACCATGGCTCCATAACTTAGGACAACATATCTATGAAACCTATGGAGATA
CTTGGACAGGAGTTGAAGCAATAATAAGAATCCTGCAACAATTACTGTTTATTCATT
TCAGGATTGGGTGCCATCATAGCAGAATAGGCATTTTGCGACAGAGAAGAGCAAGA
AATGGAGCCAATAGATCCTAACCTAGAACCCTGGAACCATCCAGGAAGTCAGCCTA
AAACTGCTTGTAATGGGTGTTACTGTAAACGTTGCAGCTATCATTGTCTAGTTTGCTT
TCAGAAAAAAGGCTTAGGCATTTACTATGGCAGGAAGAAGCGGAGACAGCGACGAA
GCGCTCCTCCAAGCAATAAAGATCATCAAGATCCTCTACCAAAGCAGTAAGTACCG
AATAGTATATGTAATGTTAGATTTAACTGCAAGAATAGATTCTAGATTAGGAATAGG
AGCATTGATAGTAGCACTAATCATAGCAATAATAGTGTGGACCATAGTATATATAG
AATATAGGAAATTGGTAAGGCAAAGGAAAATAGACTGGTTAGTTAAAAGGATTAGG
GAAAGAGCAGAAGACAGTGGCAATGAGAGCGAGGGGGATACTGAAGAATTATCGA
CACTGGTGGATATGGGGCATCTTAGGCTTTTGGATGCTAATGATGTGTAATGTGAA
GGGCTTGTGGGTCACAGTCTACTACGGGGTACCTGTGGGGAGAGAAGCAAAAACT
ACTCTATTTTGTGCATCAGATGCTAAAGCATATGAGAAAGAAGTGCATAATGTCTG
GGCTACACATGCCTGTGTACCCACAGACCCCAACCCACAAGAAGTGATTTTGGGC
AATGTAACAGAAAATTTTAACATGTGGAAAAATGACATGGTGGATCAGATGCAGG
AAGATATAATCAGTTTATGGGATCAAAGCCTTAAGCCATGTGTAAAATTGACCCCA
CTCTGTGTCACTTTAAACTGTACAAATGCAACTGTTAACTACAATAATACCTCTAAA
GACATGAAAAATTGCTCTTTCTATGTAACCACAGAATTAAGAGATAAGAAAAAGAA
AGAAAATGCACTTTTTTATAGACTTGATATAGTACCACTTAATAATAGGAAGAATGG
GAATATTAACAACTATAGATTAATAAATTGTAATACCTCAGCCATAACACAAGCCTG
TCCAAAAGTCTCGTTTGACCCAATTCCTATACATTATTGTGCTCCAGCTGGTTATGCG
CCTCTAAAATGTAATAATAAGAAATTCAATGGAATAGGACCATGCGATAATGTCAG
CACAGTACAATGTACACATGGAATTAAGCCAGTGGTATCAACTCAATTACTGTTAAA
TGGTAGCCTAGCAGAAGAAGAGATAATAATTAGATCTGAAAATCTGACAAACAATG
TCAAAACAATAATAGTACATCTTAATGAATCTATAGAGATTAAATGTACAAGACC

FIG. 16C

TGGCAATAATACAAGAAAGAGTGTGAGAATAGGACCAGGACAAGCATTCTATGCA
ACAGGAGACATAATAGGAGATATAAGACAAGCACATTGTAACATTAGTAAAAATGA
ATGGAATACAACTTTACAAAGGGTAAGTCAAAAATTACAAGAACTCTTCCCTAATA
GTACAGGGATAAAATTTGCACCACACTCAGGAGGGGACCTAGAAATTACTACACAT
AGCTTTAATTGTGGAGGAGAATTTTTCTATTGCAATACAACAGACCTGTTTAATAGT
ACATACAGTAATGGTACATGCACTAATGGTACATGCATGTCTAATAATACAGAGCG
CATCACACTCCAATGCAGAATAAAACAAATTATAAACATGTGGCAGGAGGTAGGAC
GAGCAATGTATGCCCCTCCCATTGCAGGAAACATAACATGTAGATCAAATATTACA
GGACTACTATTAACACGTGATGGAGGAGATAATAATACTGAAACAGAGACATTCAG
ACCTGGAGGAGGAGACATGAGGGACAATTGGAGAAGTGAATTATATAAATACAAG
GTGGTAGAAATTAAACCATTAGGAGTAGCACCCACTGCTGCAAAAAGGAGAGTGGT
GGAGAGAGAAAAAAGAGCAGTAGGAATAGGAGCTGTGTTCCTTGGGTTCTTGGGAG
CAGCAGGAAGCACTATGGGCGCAGCATCAATAACGCTGACGGTACAGGCCAGACAA
TTATTGTCTGGTATAGTGCAACAGCAAAGTAATTTGCTGAGGGCTATAGAGGCGCAA
CAGCATATGTTGCAACTCACGGTCTGGGGCATTAAGCAGCTCCAGGCAAGAGTCCTG
GCTATAGAGAGATACCTACAGGATCAACAGCTCCTAGGACTGTGGGGCTGCTCTGG
AAAACTCATCTGCACCACTAATGTGCTTTGGAACTCTAGTTGGAGTAATAAAACTCA
AAGTGATATTTGGGATAACATGACCTGGATGCAGTGGGATAGGGAAATTAGTAATT
ACACAAACACAATATACAGGTTGCTTGAAGACTCGCAAAGCCAGCAGGAAAGAAA
TGAAAAAGATTTACTAGCATTGGACAGGTGGAACAATCTGTGGAATTGGTTTAGCAT
AACAAATTGGCTGTGGTATATAAAAATATTCATAATGATAGTAGGAGGCTTGATAG
GTTTAAGAATAATTTTTGCTGTGCTCTCTCTAGTAAATAGAGTTAGGCAGGGATACT
CACCCTTGTCATTGCAGACCCTTATCCCAAACCCGAGGGGACCCGACAGGCTCGGA
GGAATCGAAGAAGAAGGTGGAGAGCAAGACAGCAGCAGATCCATTCGATTAGTGA
GCGGATTCTTGACACTTGCCTGGGACGACCTACGAAGCCTGTGCCTCTTCTGCTACC
ACCGATTGAGAGACTTCATATTAATTGTAGTGAGAGCAGTGGAACTTCTGGGACAC
AGTAGTCTCAGGGGACTGCAGAGGGGGTGGGGAACCCTTAAGTATTTGGGGAGTCT
TGTGCAATATTGGGGTCTAGAGTTAAAAAAGAGTGCTATTAATCTGCTTGATACTAT
AGCAATAGCAGTAGCTGAAGGAACAGATAGGATTCTAGAATTCATACAAAACCTTT
GTAGAGGTATCCGCAACGTACCTAGAAGAATAAGACAGGGCTTCGAAGCAGCTTTG
CAATAAAATGGGGGGCAAGTGGTCAAAAAGCAGTATAATTGGATGGCCTGAAGTAA
GAGAAAGAATCAGACGAACTAGGTCAGCAGCAGAGGGAGTAGGATCAGCGTCTCA
AGACTTAGAGAAACATGGGGCACTTACAACCAGCAACACAGCCCACAACAATGCTG
CTTGCGCCTGGCTGGAAGCGCAAGAGGAGGAAGGAGAAGTAGGCTTTCCAGTCAGA
CCTCAGGTACCTTTAAGACCAATGACTTATAAAGCAGCAATAGATCTCAGCTTCTTT
TTAAAAGAAAAGGGGGGACTGGAAGGGTTAATTTACTCCAAGAAAAGGCAAGAGAT
CCTTGATTTGTGGGTTTATAACACACAAGGCTTCTTCCCTGATTGGCAAAACTACAC
ACCGGGACCAGGGGTCAGATTTCCACTGACCTTTGGATGGTACTTCAAGCTAGAGCC
AGTCGATCCAAGGGAAGTAGAAGAGGCCAATGAAGGAGAAAACAACTGTTTACTAC
ACCCTATGAGCCAGCATGGAATGGAGGATGAAGACAGAGAAGTATTAAGATGGAAG
TTTGACAGTACGCTAGCACGCAGACACATGGCCCGCGAGCTACATCCGGAGTATTAC
AAAGACTGCTGACACAGAAGGGACTTTCCGCTGGGACTTTCCACTGGGGCGTTCCAG
GAGGTGTGGTCTGGGCGGGACAGGGGAGTGGTCAGCCCTGAGATGCTGCATATAAG
CAGCTGCTTTTCGCCTGTACTGGGTCTCTCTAGGTAGACCAGATCTGAGCCCGGGAG

FIG. 16D

CTCTCTGGCTATCTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTG
CCTTGAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGA
CCACTTGTGGTAGTGTGGAAAATCTCTAGCA

FIG. 16E